Optimizing convex functions over nonconvex sets

Daniel Bienstock and Alexander Michalka
Columbia University
New York

xpl.tex.Fri:Dec:09:150952:2011

Abstract

In this paper we derive strong linear inequalities for sets of the form

$$\{ (x, q) \in \mathbb{R}^d \times \mathbb{R} \ : \ q \ge Q(x), \ \ x \in \mathbb{R}^d - \mathrm{int}(P) \},$$

where $Q(x) : \mathbb{R}^d \to \mathbb{R}$ is a quadratic function, $P \subset \mathbb{R}^d$ and "int" denotes interior. Of particular but not exclusive interest is the case where P denotes a closed convex set. In this paper, we present several cases where it is possible to characterize the convex hull by efficiently separable linear inequalities.

1 The positive-definite case

We consider sets of the form

$$S \ \doteq \ \{ (x, q) \in \mathbb{R}^d \times \mathbb{R} \ : \ q \ge Q(x), \ \ x \in \mathbb{R}^d - \mathrm{int}(P) \}, \tag{1}$$

where $Q(x) : \mathbb{R}^d \to \mathbb{R}$ is a *positive-definite* quadratic function, and each connected component of $P \subset \mathbb{R}^d$ is a homeomorph of either a half-plane or a ball. Thus, each connected component of P is a closed set with nonempty interior.

Since $Q(x)$ is positive definite, we may assume without loss of generality that $Q(x) = \|x\|^2$ (achieved via a linear transformation). For any $y \in \mathbb{R}^d$, the linearization inequality

$$q \ \ge \ 2y^T(x - y) + \|y\|^2 \ = \ 2y^T x - \|y\|^2 \tag{2}$$

is valid for all $(x, q) \in \mathbb{R}^d \times \mathbb{R}$. We seek ways of making this inequality stronger.

Definition 1.1 *Given $\mu \in \mathbb{R}^d$ and $R \ge 0$, we write $\mathcal{B}(\mu, R) \ = \ \{ x \in \mathbb{R}^d \ : \ \|x - \mu\| \le R \}$.*

1.1 Geometric characterization

Let $x \in \mathbb{R}^d$. Then $x \in \mathbb{R}^d - \mathrm{int}(P)$ if and only if

$$\|x - \mu\|^2 \ge \rho, \quad \text{for each ball } \mathcal{B}(\mu, \sqrt{\rho}) \subseteq P. \tag{3}$$

In terms of our set S, we can rewrite (3) as

$$q \ \ge \ 2\mu^T x - \|\mu\|^2 + \rho, \quad \text{for each ball } \mathcal{B}(\mu, \sqrt{\rho}) \subseteq P. \tag{4}$$

On the other hand, suppose

$$\delta q \ - \ 2\beta^T x \ \ \ge \ \ \beta_0 \tag{5}$$

is valid for S. Since $\mathbb{R}^d - P$ contains points with arbitrarily large norm it follows $\delta \ge 0$. Suppose that $\delta > 0$: then without loss of generality $\delta = 1$. Further, given $x \in \mathbb{R}^d$, (5) is satisfied by (x, q) with $q \ge \|x\|^2$ if and only if it is satisfied by $(x, \|x\|^2)$, and so if and only if we have

$$\|x - \beta\|^2 \ge \ \ \|\beta\|^2 + \beta_0. \tag{6}$$

Since (5) is valid for S, we have that (6) holds for each $x \in \mathbb{R}^d - \mathrm{int}(P)$. Assuming further that (5) is not trivial, that is to say, it is violated by some $(z, \|z\|^2)$ with $z \in \mathrm{int}(P)$, we must therefore have that $\|\beta\|^2 + \beta_0 > 0$ and $\mathcal{B}(\beta, \sqrt{\|\beta\|^2 + \beta_0}) \subseteq P$, i.e. statement (6) is an example of (3). Below we discuss several ways of sharpening these observations.

1.2 Lifted first-order cuts

Let $y \in \partial P$. Then we can always find a ball $\mathcal{B}(\mu, \sqrt{\rho}) \subseteq P$ such that $\|\mu - y\|^2 = \rho$, possibly by setting $\mu = y$ and $\rho = 0$.

Definition 1.2 *Given $y \in \partial P$, we say P is locally flat at y if there is a ball $\mathcal{B}(\mu, \sqrt{\rho}) \subseteq P$ with $\|\mu - y\|^2 = \rho$ and $\rho > 0$.*

Suppose P is locally flat at y and let $\mathcal{B}(\mu, \sqrt{\rho})$ be as in the definition. Let $a^T x \geq a_0$ be a supporting hyperplane for $\mathcal{B}(\mu, \sqrt{\rho})$ at y, i.e. $a^T y = a_0$ and $a^T x \geq a_0$ for all $x \in \mathcal{B}(\mu, \sqrt{\rho})$. We claim that

$$q \;\; \geq \;\; 2y^T x \, - \, \|y\|^2 \, + \, 2\alpha(a^T x - a_0) \tag{7}$$

is valid for S if $\alpha \geq 0$ is small enough. To see this, note that since $a^T x \geq a_0$ supports $\mathcal{B}(\mu, \sqrt{\rho})$ at y, it follows that $\mu - y = \bar{\alpha} a$ for small enough, but positive $\bar{\alpha}$, i.e.,

$$\mathcal{B}(y + \bar{\alpha} a \, , \, \sqrt{\bar{\alpha}^2 \|a\|^2}) \;\; = \;\; \mathcal{B}(\mu, \sqrt{\rho}). \tag{8}$$

Now, assume $\alpha \leq \bar{\alpha}$. Then $(v, \|v\|^2)$ violates (7) iff

$$\begin{aligned}
\|v\|^2 \;\; &< \;\; 2y^T v \, - \, \|y\|^2 \, + \, 2\alpha(a^T v - a_0) && (9)\\
&= \;\; 2(y + \alpha a)^T v \, - \, \|y + \alpha a\|^2 \, + \alpha^2 \|a\|^2 + \, 2\alpha(y^T a - a_0) && (10)\\
&= \;\; 2(y + \alpha a)^T v \, - \, \|y + \alpha a\|^2 \, + \alpha^2 \|a\|^2, \quad \text{that is,} && (11)\\
v \;\; &\in \;\; \mathcal{B}(y + \alpha a \, , \, \sqrt{\alpha^2 \|a\|^2}) \, \subset \, \mathcal{B}(\mu, \sqrt{\rho}) && (12)
\end{aligned}$$

since $\alpha \leq \bar{\alpha}$. In other words, for small enough, but positive α, (7) is valid for S.

In fact, the above derivation implies a stronger statement: since $a^T x \geq a_0$ supports $\mathcal{B}(y + \alpha a \, , \sqrt{\alpha^2 \|a\|^2})$ at y, for any $\alpha > 0$, it follows (7) is valid for S iff $\mathcal{B}(y + \alpha a \, , \, \sqrt{\alpha^2 \|a\|^2}) \subseteq P$. Define

$$\hat{\alpha} \;\doteq\; \sup\{ \, \alpha \, : \, (7) \text{ is valid} \, \}.$$

If there exists $v \notin P$ such that $a^T v > a_0$ then the assumptions on P imply that $\hat{\alpha} < +\infty$ and the 'sup' is a 'max'. If on the other hand $a^T v \leq a_0$ for all $v \notin P$ then $\hat{\alpha} = +\infty$ (and, of course, $a^T x \leq a_0$ is valid for S). In the former case, we call

$$q \;\; \geq \;\; 2y^T x \, - \, \|y\|^2 \, + \, 2\hat{\alpha}(a^T x - a_0) \tag{13}$$

a *lifted first-order inequality.*

Theorem 1.3 *Any linear inequality*

$$\delta q \; - \; \beta^T x \;\; \geq \;\; \beta_0 \tag{14}$$

valid for S either has $\delta = 0$ (in which case the inequality is valid for $\mathbb{R}^d - P$), or $\delta > 0$ and (14) is dominated by a lifted first-order inequality or by a linearization inequality (2).

Proof. Consider a valid inequality (14). As above we either have $\delta = 0$, in which case we are done, or without loss of generality $\delta = 1$, and by increasing β_0 if necessary we have that (14) is tight at some point $(y, \|y\|^2) \in \mathbb{R}^d \times \mathbb{R}$.
Write

$$\beta^T x + \beta_0 \;\; = \;\; 2y^T x \, - \, \|y\|^2 \, + \, 2\gamma^T x \, + \, \gamma_0, \tag{15}$$

for appropriate γ and γ_0. Suppose first that $y \in \text{int}(\mathbb{R}^d - P)$. Then $(\gamma, \gamma_0) = (0, 0)$, or else (14) would not be valid in a neighborhood of y. Thus, (14) is a linearization inequality.

Suppose next that $y \in \partial P$, and that (14) is not a linearization inequality, i.e. $(\gamma, \gamma_0) \neq (0, 0)$. We can write (14) as

$$\begin{aligned} q &\geq 2y^T x - \|y\|^2 + 2\gamma^T x + \gamma_0 \\ &= 2(y+\gamma)^T x - \|y+\gamma\|^2 - 2\gamma^T y - \|\gamma\|^2 + \gamma_0. \end{aligned} \tag{16}$$

Since (14) is not a linearization inequality, and is satisfied at $(y, \|y\|^2)$ there exist points $(v, \|v\|^2)$ (with v near y) which do not satisfy it. Necessarily, any such v must not lie in $\mathbb{R}^d - P$ (since (14) is valid for S). Using (16) this happens iff

$$\|v\|^2 < 2(y+\gamma)^T v - \|y+\gamma\|^2 - 2\gamma^T y - \|\gamma\|^2 + \gamma_0, \quad \text{that is,} \tag{17}$$

$$v \in \operatorname{int}\left(\mathcal{B}\left(y+\gamma\,,\, \sqrt{-2\gamma^T y - \|\gamma\|^2 + \gamma_0} \right) \right). \tag{18}$$

In other words, the set of points that violate (14) is the interior of some ball $\mathcal{B}$ with positive radius, which necessarily must be contained in P. Since $(y, \|y\|^2)$ satisfies (14) with inequality, y is in the boundary of $\mathcal{B}$. Thus, P is locally flat at y; writing $a^T x = a_0$ to denote the hyperplane orthogonal to γ through y, we have that (14) is dominated by the resulting lifted first-order inequality. ■

1.3 The polyhedral case

Here we will discuss an efficient separation procedure for lifted first-order inequalities in the case that P is a polyhedron. Further properties of these inequalities are discussed in [10].

Suppose that $P = \{ x \in \mathbb{R}^d \,:\, a_i^T x \geq b_i,\ 1 \leq i \leq m \}$ is a full-dimensional polyhedron, where each inequality is facet-defining and the representation of P is minimal. For $1 \leq i \leq m$ let $H_i \doteq \{ x \in \mathbb{R}^d \,:\, a_i^T x = b_i \}$. For $i \neq j$ let $H_{\{i,j\}} \doteq \{ x \in \mathbb{R}^d \,:\, a_i^T x = b_i,\, a_j^T x = b_j \}$. $H_{\{i,j\}}$ is $(d-2)-$dimensional; we denote by ω_{ij} the unique unit norm vector orthogonal to both H_{ij} and a_i (unique up to reversal).

Consider a fixed pair of indices $i \neq j$, and let $\mu \in \operatorname{int}(P)$. Let Ω_{ij} be the 2-dimensional hyperplane through μ generated by a_i and ω_{ij}. By construction, therefore, Ω_{ij} is orthogonal to $H_{\{i,j\}}$ and is thus the orthogonal complement to $H_{\{i,j\}}$ through μ. It follows that $\Omega_{ij} = \Omega_{ji}$ and that this hyperplane contains the orthogonal projection of μ onto H_i (which we denote by $\pi_i(\mu)$ and the orthogonal projection of μ onto H_j ($\pi_j(\mu)$, respectively). Further, $\Omega_{ij} \cap H_{\{i,j\}}$ consists of a single point $k_{\{i,j\}}(\mu)$ satisfying

$$\begin{aligned} \|\mu - k_{\{i,j\}}(\mu)\|^2 &= \|\mu - \pi_i(\mu)\|^2 + \|\pi_i(\mu) - k_{\{i,j\}}(\mu)\|^2 \\ &= \|\mu - \pi_j(\mu)\|^2 + \|\pi_j(\mu) - k_{\{i,j\}}(\mu)\|^2. \end{aligned} \tag{19}$$

Now we return to the question of separating lifted first-order inequalities. Note that P is locally flat at a point y if and only if y is in the relative interior of one of the facets. Suppose that y is in the relative interior of the i^{th} facet. Then the lifting coefficient corresponding to the lifted first-order inequality at y is tight at some other point $\hat{y}$ in a different facet, facet j, say. Thus, there is a ball $\mathcal{B}(\mu, \sqrt{\rho})$ contained in P which is tangent to H_i at y and tangent to H_j at $\hat{y}$, that is to say,

$$y = \pi_i(\mu) \text{ and } \hat{y} = \pi_j(\mu), \tag{20}$$

$$y - k_{\{i,j\}}(\mu) \text{ is parallel to } \omega_{ij} \text{ and } \hat{y} - k_{\{i,j\}}(\mu) \text{ is parallel to } \omega_{ji}, \tag{21}$$

$$\|\mu - y\|^2 = \|\mu - \hat{y}\|^2 = \rho, \quad \text{and by (19),} \tag{22}$$

$$\|y - k_{\{i,j\}}(\mu)\| = \|\hat{y} - k_{\{i,j\}}(\mu)\|, \quad \text{and} \tag{23}$$

$$\|\mu - y\| = \tan\phi\, \|y - k_{\{i,j\}}(\mu)\|, \tag{24}$$

where 2ϕ is the angle formed by ω_{ij} and ω_{ji}. By the preceding discussion, $\rho = \hat{\alpha}^2 \|a_i\|^2$; using (22) and (24) we will next argue that the lifting coefficient, $\hat{\alpha}$, is an **affine** function of y.

Let $h^g_{\{i,j\}}$ $(1 \leq g \leq d-2)$ be a basis for $\{ x \in \mathbb{R}^d \,:\, a_i^T x = a_j^T x = 0 \}$. Then a_i, together with ω_{ij} and the $h^g_{\{i,j\}}$ form a basis for $\mathbb{R}^d$. Let

- O_i be the projection of the origin onto H_i – hence O_i is a multiple of a_i,

- N_i be the projection of O_i onto $H_{\{i,j\}}$.

We have

$$y \;=\; O_i \,+\, (N_i - O_i) \,+\, (k_{\{i,j\}}(\mu) - N_i) \,+\, (y - k_{\{i,j\}}(\mu)), \tag{25}$$

and thus, since $N_i - O_i$ and $y - k_{\{i,j\}}(\mu)$ are parallel to ω_{ij}, and $k_{\{i,j\}}(\mu) - N_i$ and O_i are orthogonal to ω_{ij},

$$\omega_{ij}^T y \;=\; \omega_{ij}^T(N_i - O_i) + \omega_{ij}^T(y - k_{\{i,j\}}(\mu)) \;=\; \omega_{ij}^T(N_i - O_i) + \|\omega_{ij}\| \|y - k_{\{i,j\}}(\mu)\|, \tag{26}$$

or

$$\|y - k_{\{i,j\}}(\mu)\| \;=\; \|\omega_{ij}\|^{-1} \omega_{ij}^T\,(y - N_i + O_i). \tag{27}$$

Consequently,

$$\hat{\alpha} \;=\; \frac{\rho}{\|a_i\|} \;=\; \frac{\tan\phi}{\|a_i\|}\,\|y - k_{\{i,j\}}(\mu)\| \tag{28}$$

$$\;=\; \frac{\tan\phi}{\|a_i\|}\,\|\omega_{ij}\|^{-1} \omega_{ij}^T\,(y - N_i + O_i). \tag{29}$$

We will abbreviate this expression as $p_{ij} y + q_{ij}$. Let $x^* \in \mathbb{R}^d$. The problem of finding the strongest possible lifted first-order inequality at x^* chosen from among those obtained by starting from a point on face i, can be written as follows:

$$\min \quad -2y^T x^* \,+\, \|y\|^2 \,-\, 2\alpha(a^T x^* - a_0) \tag{30}$$

$$s.t. \quad y \in P \tag{31}$$

$$a_i^T y = b_i \tag{32}$$

$$0 \le \alpha \le p_{ij} y + q_{ij} \quad \forall\; j \ne i. \tag{33}$$

This is a linearly constrained, convex quadratic program with $d+1$ variables and $2m-1$ constraints. By solving this problem for each choice of $1 \le i \le m$ we obtain the the strongest inequality overall.

1.3.1 The Disjunctive Approach

For $1 \le i \le m$ let $\bar{P}^i = \{\, x \in \mathbb{R}^d \,:\, a_i^T x \le b_i \}$; thus $\mathbb{R}^d - P = \bigcup_i \bar{P}^i$. Further, for $1 \le i \le m$ write:

$$\bar{Q}^i = \{\, (x, q) \in \mathbb{R}^d \times R \,:\, a_i^T x \le b_i,\; q \ge \|x\|^2 \}.$$

Thus, $(x^*, q^*) \in \mathrm{conv}(S)$ if and only if (x^*, q^*) can be written as a convex combination of points in the sets $\bar{Q}^i$. This is the approach pioneered in Ceria and Soares [6] (also see [13]). The resulting separation problem is carried out by solving a second-order cone program with m conic constraints and md variables, and then using second-order cone duality in order to obtain a linear inequality (details in [10]).

Thus, the derivation we presented above amounts to a possibly simpler alternative to the Ceria-Soares approach, which also makes explicit the geometric nature of the resulting cuts.

1.4 The ellipsoidal case

In this section we will discuss an efficient separation procedure for lifted first-order inequalities in the case that P is a convex ellipsoid, in other words,

$$P \,=\, \{x \,\in \mathbb{R}^d \,:\, x^T A x \,-\, 2c^T x \,+\, b \,\le\, 0\},$$

for appropriate $A \succeq 0$, c and b. The separation problem to solve can be written as follows: given$(x^*, q^*) \in \mathbb{R}^{d+1}$,

$$\max\{\, \Theta(\rho) \,:\, \rho \ge 0 \,\} \;\; \text{where, for fixed } \rho \ge 0, \tag{34}$$

$$\Theta(\rho) \doteq \max\ \rho - (q^* - 2\mu^T x^* + \mu^T\mu) = \rho - \|x^* - \mu\|^2 - q^* + \|x^*\|^2 \quad (35)$$
$$\text{s.t. } \mu \in \mathbb{R}^d,\ \rho \geq 0 \text{ and } \mathcal{B}(\mu, \sqrt{\rho}) \subseteq P \quad (36)$$

Consider a **fixed** value $\rho > 0$. We will first show that with this proviso the condition

$$\mathcal{B}(\mu, \sqrt{\rho}) \subseteq P \quad (37)$$

is SOCP-representable. We note that [1] considers the problem of finding a minimum-radius ball containing a family of ellipsoids; our separation problem addresses, in a sense, the opposite situation, which leads to a somewhat different analysis. Our equations (40)-(41)are related to formulae found in [1] (also see [7]) but again reflecting the opposite nature of the problem. Also see [4]. Some of the earliest studies in this direction are found in [8].

Returning to (37), notice that this condition is equivalent to stating

$$\|x\|^2 - \rho \geq 0, \qquad \forall\, x \ \text{ s.t. } (x+\mu)^T A(x+\mu) - 2c^T(x+\mu) + b \geq 0. \quad (38)$$

Using the S-Lemma [14], [11], [2], (38) holds if and only if there exists a quantity $\theta \geq 0$ such that, for **all** $x \in \mathbb{R}^d$

$$x^T(I - \theta A)\,x + 2\theta(c^T - \mu^T A)x + \theta(-\mu^T A\mu + 2c^T\mu - b) - \rho \geq 0.$$

Clearly we must have $\theta > 0$; writing $\tau = \theta^{-1}$ we have that (38) holds if and only if there exists $\tau > 0$ such that

$$x^T\left(I - \frac{1}{\tau}A\right)x + \frac{2}{\tau}(c^T - \mu^T A)x + \frac{1}{\tau}(-\mu^T A\mu + 2c^T\mu - b) - \rho \geq 0 \ \ \forall x \in \mathbb{R}^d. \quad (39)$$

Let the eigenspace decomposition of A be $A = U\Lambda U^T$ and write

$$y \doteq U^T x, \text{ and } v = v(\mu) \doteq U^T(c - A\mu).$$

Then we have that (39) holds iff for all $y \in \mathbb{R}^d$,

$$y^T\left(I - \frac{1}{\tau}\Lambda\right)y + \frac{2}{\tau}v^T y + \frac{1}{\tau}(-\mu^T A\mu + 2c^T\mu - b) - \rho \geq 0,$$

or, equivalently,

$$\begin{pmatrix} I - \frac{1}{\tau}\Lambda & \frac{1}{\tau}v \\ & \\ \frac{1}{\tau}v^T & \frac{1}{\tau}(-\mu^T A\mu + 2c^T\mu - b) - \rho \end{pmatrix} \succeq 0. \quad (40)$$

Let λ_{max} denote the largest eigenvalue of A. Then (40) holds iff $\tau \geq \lambda_{max}$, and

$$-\frac{1}{\tau^2}\sum_{j=1}^{d}\frac{v_j^2}{1 - \lambda_j/\tau} + \frac{1}{\tau}(-\mu^T A\mu + 2c^T\mu - b) - \rho \geq 0,$$

or, equivalently

$$-\sum_{j=1}^{d}\frac{v_j^2}{\tau - \lambda_j} - \mu^T A\mu + 2c^T\mu - b - \rho\tau \geq 0 \quad (41)$$

which is SOCP-representable. Formally this is done as follows: (41) holds iff there exist quantities y_j, $1 \leq j \leq d$ such that

$$y_j(\tau - \lambda_j) \ \geq \ v_j^2, \ \ 1 \leq j \leq d, \text{ and } -\sum_{j=1}^{d} y_j - \mu^T A\mu + 2c^T\mu - b - \rho\tau \geq 0. \quad (42)$$

In summary, then, for fixed ρ the problem of finding the most violated lifted first-order inequality can be formulated as the following SOCP, with variables μ, τ, v and y:

$$\begin{aligned}
\min \quad & 2\mu^T x^* + \mu^T \mu \ + q^* - \rho & (43)\\
\text{s.t.} \quad & v = U^T(c - A\mu) & (44)\\
& \tau \geq \lambda_{max} & (45)\\
& y_j(\tau - \lambda_j) \ \geq \ v_j^2, \ \ 1 \leq j \leq d, & (46)\\
& -\sum_{j=1}^{d} y_j \ + 2c^T\mu \ - \ b \ - \ \rho\tau \ \geq \ \mu^T A \mu. & (47)
\end{aligned}$$

Here, constraints (46) and (47) are conic (in (47), it is critical that ρ is a fixed value, since τ is a variable).

Lemma 1.4 *Let K be an arbitrary convex set and $v \in K$. For $\rho > 0$ the function*

$$\begin{aligned}
N(\rho) \quad & \doteq \quad \min \|v - \mu\|^2 \\
s.t. \quad & \quad \mathcal{B}(\mu, \sqrt{\rho}) \subseteq K, & (48)
\end{aligned}$$

is convex.

Pending the proof of this result, we note that as per eq. (35), if $A \succeq 0$ then $\Theta(\rho)$ is a concave function of ρ. Thus the separation problem can solved to arbitrary tolerance using e.g. golden ratio search, with the SOCP (43)-(47) as a subroutine.

Proof of Lemma 1.4. To prove convexity of N, it suffices to show that for any pair of values $\rho_1 \neq \rho_2$ there exists a function $g(\rho)$ such that

(a) $g(\rho_i) = N(\rho_i), \ i = 1, 2,$

(b) $g(\rho) \geq N(\rho_i)$ for every ρ between ρ_1 and ρ_2,

(c) $g(\rho)$ is convex between ρ_1 and ρ_2.

Thus, let ρ_1, ρ_2 be given. For $i = 1, 2$ let $\mu_i = \text{argmin}\, N(\rho_i)$ and $R_i = \sqrt{\rho_i}$. Assume without loss of generality that $R_1 < R_2$. Let $0 \leq \lambda \leq 1$. Since K is convex,

$$\mathcal{B}(\ (1 - \lambda)\mu_1 + \lambda\mu_2 \ , \ \sqrt{((1 - \lambda)R_1 + \lambda R_2)^2} \ ) \quad \subseteq \quad K, \tag{49}$$

in other words, for any point μ in the segment $[\mu_1, \mu_2]$, there is a ball with center μ, contained in K and with radius

$$R_1 + \frac{R_2 - R_1}{\|\mu_2 - \mu_1\|}\|\mu - \mu_1\|, \tag{50}$$

or, to put it even more explicitly, as a point μ moves from μ_1 to μ_2 there is a ball with center μ contained in K, whose radius is obtained by linearly interpolating between R_1 and R_2 . Let μ^* be the nearest point to v on the line defined by μ_1 and μ_2 (possibly $\mu^* \notin K$). For $i = 1, 2$, let $t_i \doteq \|\mu^* - \mu_i\|$.

Suppose first that μ^* is in the line segment between μ_1 and μ_2 and $\mu^* \neq \mu_1$. By (49) there is a ball centered at μ and contained in K with radius strictly larger than R_1, a contradiction by definition of μ_1. The same contradiction would arise if μ_2 separates μ^* and μ_1.

Thus μ_1 separates μ^* and μ_2. Defining

$$s \quad = \quad \frac{R_2 - R_1}{t_2 - t_1} \ > \ 0, \tag{51}$$

we have that for $-t_1 \leq t \leq t_2 - 2t_1$ the point

$$\mu(t) \quad = \quad \mu_1 + \frac{\mu_2 - \mu_1}{t_2 - t_1}(t + t_1) \tag{52}$$

lies in the segment $[\mu_1, \mu_2]$ and is the center of a ball of radius

$$R(t) \quad = \quad R_1 \; + \; s(\, t_1 \, + \, t\,); \tag{53}$$

further $\mu(-2t_1) = \mu^*$. Since K is convex, the segment between v and μ_2 is contained in K; let w be the point in that segment with $\|v - w\| = \|v - \mu_1\|$; by the triangle inequality

$$\|\mu_1 - \mu_2\| \geq \|w - \mu_2\|. \tag{54}$$

Let π be the slope of the linear interpolant, between values R^* and R_2, along the segment $[v, \mu_2]$, i.e. $R^* + \pi\|v - \mu_2\| = R_2$. Then, as previously, $\mathcal{B}(w, \sqrt{R_w}) \subseteq K$ where $R_w = R^* + \pi\|v - w\|$. But then it follows by definition of μ_1 that

$$R_w \leq R_1. \tag{55}$$

By (55) and (54), we have $\pi \geq s$, and therefore, by (55),

$$R_1 \;\geq\; R^* \;+\; \pi\|v - w\| \;=\; R^* \;+\; \pi\|v - \mu_1\| \;\geq\; R^* \;+\; \pi\|\mu^* - \mu_1\| \;\geq\; st_1, \tag{56}$$

Now, for any t, since $\mu(-2t_1) = \mu^*$,

$$\|\mu(t) - \mu^*\| \quad = \quad \frac{\|\mu_2 - \mu_1\|}{t_2 - t_1}|t + 2t_1|. \tag{57}$$

Define $\gamma = (\mu_2 - \mu_1)/(t_2 - t_1)$,and

$$g(\rho) \quad \doteq \quad \gamma^2 \left(\frac{\sqrt{\rho} - R_1}{s} \; + t_1 \right)^2 \quad + \quad \|\mu^* - v\|^2.$$

We will prove that g satisfies properties (i)-(iii) listed above. For ρ between ρ_1 and ρ_2, writing $R = \sqrt{\rho}$ and

$$t = (R - R_1)/s - t_1,$$

it follows that $\mu(t)$ is the center of a ball of radius R contained in K. Further, since $\|\mu(t) - \mu^*\| = \gamma|t + 2t_1|$,

$$g(\rho) \;=\; \|\mu(t) - \mu^*\|^2 \quad + \quad \|\mu^* - v\|^2, \tag{58}$$

and so g satisfies (i) and (ii). Finally, to see that g is convex, note that the coefficient of $\sqrt{\rho}$ in the expansion of $g(\rho)$ in (58) equals

$$2\gamma^2 \left(\frac{t_1}{s} \; - \; \frac{R_1}{s^2} \right) \;\leq\; 0, \tag{59}$$

by (56). ■

Note: We speculate that $A \succeq 0$ (i.e., convexity of P) is not required for Lemma 1.4, and that further the overall separation algorithm can be improved to avoid dealing with the fixed ρ case.

2 Indefinite Quadratics

The general case of a set $\{\,(x, q) \in \mathbb{R}^d \times \mathbb{R} \;:\; q \geq Q(x), \;\; x \in \mathbb{R}^d - \mathrm{int}(P)\,\}$, where $Q(x)$ is a semidefinite quadratic can be approached in much the same way as that employed above, but with some important differences.

We first consider the case where P is a polyhedron. Let $P = \{(x, w) \in \mathbb{R}^{d+1} \;:\; a_i^T x - w \leq b_i, \; 1 \leq i \leq m\}$ (here, w is a scalar). Consider a set of the form

$$S \quad \doteq \quad \left\{\, (x, w, q) \in \mathbb{R}^{d+2} \;:\; q \geq \|x\|^2, \; (x, w) \in \mathbb{R}^{d+1} - P \right\}. \tag{60}$$

Many examples can be brought into this form, or similar, by an appropriate affine transformation. Consider a point (x^*, w^*) in the relative interior of the i^{th} facet of P. We seek a lifted first-order inequality of the form

$$(2x^* - \alpha a_i)^T x + \alpha w + \alpha b_i - \|x^*\|^2 \leq q,$$

for appropriate $\alpha \geq 0$. If we are lifting to the j^{th} facet, then we must have $v_{ij} = \alpha b_i - \|x^*\|^2$, where

$$\begin{aligned} v_{ij} \quad &\doteq \quad \min \|x\|^2 - (2x^* - \alpha a_i)^T x - \alpha w && (61) \\ & \qquad \text{s.t.} \quad a_j^T x - w = b_j. && (62) \end{aligned}$$

To solve this optimization problem, consider its Lagrangian:

$$\mathcal{L}(x, w, \nu) = \|x\|^2 - (2x^* - \alpha a_i)^T x - \alpha w - \nu(a_j^T x - w - b_j)$$

Taking the gradient in x and setting it to 0:

$$\begin{aligned} \nabla_x \mathcal{L} = 0 \quad &\Leftrightarrow \quad 2x - 2x^* + \alpha a_i - \nu a_j = 0 \\ &\Leftrightarrow \quad x = x^* - \frac{\alpha}{2} a_i + \frac{\nu}{2} a_j \end{aligned}$$

Now doing the same for w:

$$\begin{aligned} \nabla_w \mathcal{L} = 0 \quad &\Leftrightarrow \quad -\alpha + \nu = 0 \\ &\Leftrightarrow \quad \nu = \alpha \end{aligned}$$

Combining these two gives

$$x = x^* - \frac{\alpha}{2} a_i + \frac{\alpha}{2} a_j$$

then using the constraint $a_j^T x - w = b_j$ gives

$$w = a_j^T x^* - b_j - \frac{\alpha}{2} a_j^T a_i + \frac{\alpha}{2} a_j^T a_j$$

Next we expand out the objective value using the expressions we have derived for x and w, and set the result equal to $\alpha b_i - \|x^*\|^2$. Omitting the intermediate algebra, the result is the quadratic equation

$$\alpha(a_i^T x^* - b_i - (a_j^T x^* - b_j)) - \frac{1}{4}\alpha^2 (a_i^T a_i - 2a_i^T a_j + a_j^T a_j) = 0$$

One root of this equation is $\alpha = 0$. The other root is

$$\hat{\alpha} \quad \doteq \quad \frac{4(a_i^T x^* - b_i - (a_j^T x^* - b_j))}{a_i^T a_i - 2a_i^T a_j + a_j^T a_j}. \qquad (63)$$

Since $a_i^T x^* - w^* = b_i$, and $a_j^T x^* - w^* \leq b_j$, we have

$$a_i^T x^* - b_i - (a_j^T x^* - b_j) > 0$$

so $\hat{\alpha} > 0$ (the denominator is a squared distance between some two vectors so it is non-negative). Moreover, the expression for $\hat{\alpha}$ is an affine function of x^*. Thus, as in Section 1.3, the computation of a maximally violated lifted first-order inequality is a convex optimization problem.

In this case there is an additional detail of interest: note that the points cut-off by the inequality are precisely those of the form $(x, w, \|x\|^2)$ such that

$$(2x^* - \hat{\alpha} a_i)^T x + \alpha w + \alpha b_i - \|x^*\|^2 > \|x\|^2. \qquad (64)$$

This condition defines the interior of a *paraboloid*; this is the proper generalization of condition (3) in the indefinite case.

References

[1] F. Alizadeh and D. Goldfarb, Second-Order Cone Programming, *Mathematical Programming* **95** (2001), 3 – 51.

[2] A. Ben-Tal and A. Nemirovsky, em Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications (2001) MPS-SIAM Series on Optimization, SIAM, Philadelphia, PA.

[3] P. Belotti, A.J. Miller and M. Namazifar, Valid Inequalities and Convex Hulls for Multilinear Functions, *Electronic Notes in Discrete Mathematics* **36** (2010), 805 - 812.

[4] D. Bienstock, Eigenvalue techniques for proving bounds for convex objective, nonconvex programs, *Proc. IPCO 2010*.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004).

[6] S. Ceria and J. Soares, Convex programming for disjunctive convex optimization, *Mathematical Programming* **86** (1999), 595 - 614.

[7] D. Goldfarb and G. Iyengar, Robust Portfolio Selection Problems, *Mathematics of Operations Research*, **28** (2002), 1 – 38.

[8] G.H. Golub, Some modified matrix eigenvalue problems, *SIAM Review* **15** (1973), 318 – 334.

[9] M. Kilinc, J. Linderoth and J. Luedtke, Effective Separation of Disjunctive Cuts for Convex Mixed Integer Nonlinear Programs, *Optimization Online* (2010).

[10] A. Michalka, Ph.D. Dissertation, Columbia University (in preparation).

[11] I. Pólik and T. Terlaky, A survey of the S-lemma, *SIAM Review* **49** (2007), 371 – 418.

[12] A. Qualizza, P. Belotti, F. Margot, Linear Programming Relaxations of Quadratically Constrained Quadratic Programs, manuscript, 2011.

[13] R.A. Stubbs and S. Mehrotra, A branch-and-cut method for 0-1 mixed convex programming, *Mathematical Programming* **86** (1999), 515 - 532.

[14] V. A. Yakubovich, S-procedure in nonlinear control theory, Vestnik Leningrad University, **1** (1971) 62–777.